

Lange, Sarah

Methodische Reflexionen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien

ZEP : Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik 38 (2015) 4, S. 4-12



Quellenangabe/ Reference:

Lange, Sarah: Methodische Reflexionen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien - In: ZEP : Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik 38 (2015) 4, S. 4-12 - URN: urn:nbn:de:0111-pedocs-140759 - DOI: 10.25656/01:14075

<https://nbn-resolving.org/urn:nbn:de:0111-pedocs-140759>

<https://doi.org/10.25656/01:14075>

in Kooperation mit / in cooperation with:

ZEP

Zeitschrift für internationale Bildungsforschung
und Entwicklungspädagogik

"Gesellschaft für interkulturelle Bildungsforschung und Entwicklungspädagogik e.V."

<http://www.uni-bamberg.de/allgpaed/zep-zeitschrift-fuer-internationale-bildungsforschung-und-entwicklungspaedagogik/profil>

Nutzungsbedingungen

Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Die Nutzung stellt keine Übertragung des Eigentumsrechts an diesem Dokument dar und gilt vorbehaltlich der folgenden Einschränkungen: Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use

We grant a non-exclusive, non-transferable, individual and limited right to using this document.

This document is solely intended for your personal, non-commercial use. Use of this document does not include any transfer of property rights and it is conditional to the following limitations: All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Kontakt / Contact:

peDOCS

DIPF | Leibniz-Institut für Bildungsforschung und Bildungsinformation

Informationszentrum (IZ) Bildung

E-Mail: pedocs@dipf.de

Internet: www.pedocs.de

Mitglied der


Leibniz-Gemeinschaft

ZEP

Zeitschrift für internationale Bildungsforschung
und Entwicklungspädagogik

Mit: Mitteilungen der DGfE-Kommission
Vergleichende und Internationale
Erziehungswissenschaft

4'15

Bildungsmonitoring

- Methodische Reflexionen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien
- Evidenzbasierte Entwicklung von Bildungserträgen durch Assessments
- PISA for Development
- PIAAC: Kompetenzfeststellung und -debatte in Spanien



WAXMANN

In den Diskursen um Bildungsmonitoring wird seit der Jahrtausendwende vermehrt über Leistungsfähigkeit von Bildungssystemen diskutiert. Über die regelmäßige Messung von Kompetenzen und Bildungserträgen soll insbesondere bildungspolitischen Akteuren evidenzbasiertes Steuerungswissen an die Hand gegeben werden, das als Grundlage für Entscheidungsprozesse und Bildungsreformen genutzt werden kann. Damit wird darauf abgezielt, die Qualität von Bildungssystemen zu verbessern und zu sichern. Dass solche internationalen Vergleichsstudien zunehmend auch für Länder mit mittleren Einkommensstrukturen und für Länder der Entwicklungszusammenarbeit relevant werden, zeigt deren wachsende Teilnahmefrequenz bei den letzten sechs PISA-Zyklen ebenso wie die Konzeption der OECD-Studie *PISA for Development*. Mit solchen internationalen Perspektiven auf Bildungsmonitoring sind für die jeweiligen Länder unterschiedliche Herausforderungen und Chancen verknüpft. Im Kontext dieser zunehmenden Bedeutung von international angelegten Assessments stellt sich insbesondere die Frage, wie für Staaten mit sehr divergierenden Voraussetzungen Bildungsmonitoring gestaltet werden kann.

Vor diesem Hintergrund befasst sich das Heft mit der Reflexion international vergleichender Studien – wie dem *Programme for the International Assessment of Adult Competencies* (PIAAC) und *PISA for Development* – sowie mit konstruktiven Überlegungen für ein länder- und kultursensibles Bildungsmonitoring in Ländern der Entwicklungszusammenar-

beit. Damit ist das Ziel verbunden, den Blick auf die Bedeutung von Vergleichsstudien in Ländern der Entwicklungszusammenarbeit zu richten und nach den spezifischen Herausforderungen zu fragen, wenn internationale Vergleichsstudien in Ländern durchgeführt werden, in denen bisher keine oder nur begrenzte Erfahrungen mit Assessment-Strukturen vorliegen. Damit wird das Ziel verfolgt, der interessierten Leserschaft einen fundierten Überblick in die aktuellen Diskussionen zu Assessment-Studien, die im Sinne eines längerfristigen datenbasierten Bildungsmonitorings zunehmend auch in Ländern der Entwicklungszusammenarbeit durchgeführt werden, zu geben.

Das vorliegende Heft umfasst daher die folgenden vier Themenbeiträge.

Sarah Lange zeigt mit ihrem Artikel methodische Herausforderungen auf, die mit der Teilnahme von Ländern der Entwicklungszusammenarbeit an Vergleichsstudien verbunden sind. Ausgangspunkt hierfür ist die Analyse der bisherigen Teilnahme von Ländern der Entwicklungszusammenarbeit an TIMSS und PISA. Auch die Rolle von regionalen und nationalen im Vergleich zu internationalen Vergleichsstudien wird diskutiert.

Annette Scheunpflug, Petra Stanat, Felicitas Thiel, Harm Kuper und Bettina Hannover diskutieren die Relevanz und die Möglichkeiten des Ausbaus einer Evidenzbasierung durch Assessment-Studien in Ländern der Entwicklungszusammenarbeit. Im Kontext der Überlegungen zu einem ‚länder- und kultursensiblen Monitoring‘ werden Qualitätsin-

dikatoren für Assessments vorgelegt, die potenziell zu einer Verbesserung von Bildungsqualität beitragen können.

In ihrer leitenden Position für die neue Studie „PISA for Development“ beschreiben *Michael Ward* und *Pablo Zoido* aus der Perspektive der OECD Details zur Konzeption und zur geplanten Durchführung der PISA-Studie in Ländern der Entwicklungszusammenarbeit. Dabei werden die Besonderheiten der geplanten Instrumentierung (z.B. die Haushaltserhebung von Kindern, die keine Schule besuchen) ausführlich dargestellt.

Rudolf Tippelt beschäftigt sich in seinem Beitrag mit der Frage, inwiefern internationale Vergleichsstudien, die sich mit den Kompetenzen der erwachsenen Bevölkerung auseinandersetzen, zum Monitoring von Bildungssystemen beitragen. Dazu werden empirische Ergebnisse zur Rezeption der PIAAC-Ergebnisse von weniger stark abscheidenden Ländern in der Bildungspolitik am Beispiel Spanien diskutiert.

Zudem umfasst die vorliegende Ausgabe Anmerkungen zur Neuauflage des Orientierungsrahmens für den Lernbereich Globale Entwicklung, Rezensionen sowie Informationen zum Globalen Lernen und der internationalen Bildungsforschung.

Neue Erkenntnisse und Anregungen für eine Weiterbeschäftigung mit dem Thema wünschen

Sarah Lange, Rudolf Tippelt, Julia Franz

Bamberg/München/Tübingen Januar 2016

Impressum

ZEP – Zeitschrift für internationale Bildungsforschung und Entwicklungspädagogik
ISSN 1434-4688

Herausgeber:

Gesellschaft für interkulturelle Bildungsforschung und Entwicklungspädagogik e.V. und KommEnt

Schriftleitung: Annette Scheunpflug/
Claudia Bergmüller

Redaktionsanschrift:

ZEP-Redaktion, Lehrstuhl Allgemeine Pädagogik, Markusplatz 3, 96047 Bamberg

Verlag:

Waxmann Verlag GmbH, Steinfurter Straße 555,
48159 Münster, Tel.: 0251/26 50 40
E-Mail: info@waxmann.com

Redaktion:

Barbara Asbrand, Claudia Bergmüller, Hans Bühler, Asit Datta, Julia Franz, Norbert Frieters-Reermann, Heidi Grobbauer (Österreich), Helmuth Hartmeyer (Österreich), Susanne Höck, Karola Hoffmann, Ulrich Klemm, Gregor Lang-Wojtasik, Sarah Lange, Volker Lenhart, Claudia Lohrenscheit, Bernd Overwien, Marco Rieckmann, Annette Scheunpflug, Birgit Schößwender, Klaus Seitz, Susanne Timm, Rudolf Tippelt

Technische Redaktion:

Sabine Lang (verantwortlich) 0951/863-1832, Sarah Lange (Rezensionen), Markus Ziebarth (Infos)

Anzeigenverwaltung: Waxmann Verlag GmbH, Martina Kaluza: kaluza@waxmann.com

Abbildungen: (Falls nicht bezeichnet) Privatfotos oder Illustrationen der Autoren

Titelbild: © Pixel Embargo, www.fotolia.com

Erscheinungsweise und Bezugsbedingungen: erscheint vierteljährlich; Jahresabonnement EUR 20,-, Einzelheft EUR 6,50; alle Preise verstehen sich zuzüglich Versandkosten; zu beziehen durch alle Buchhandlungen oder direkt vom Verlag. Abbestellungen spätestens acht Wochen vor Ablauf des Jahres. Das Heft ist auf umweltfreundlichem chlorfreien Papier gedruckt. Diese Publikation ist gefördert von Brot für die Welt – Evangelischen Entwicklungsdienst, Referat für Inlandsförderung, Berlin.

ZEP

Zeitschrift für internationale Bildungsforschung
und Entwicklungspädagogik

Mit: Mitteilungen der DGfE-Kommission
Vergleichende und Internationale
Erziehungswissenschaft

4'15

Themen	4	Sarah Lange Methodische Reflexionen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien
	13	Annette Scheunpflug/Petra Stanat/Felicitas Thiel/ Harm Kuper/Bettina Hannover Evidenzbasierte Entwicklung von Bildungserträgen durch Assessments – Überlegungen zu einem länder- und kultursensiblen Monitoring in Ländern der Entwicklungszusammenarbeit
	21	Michael Ward/Pablo Zoido PISA for Development
	26	Rudolf Tippelt PIAAC: Kompetenzfeststellung und -debatte in Spanien
Kommentar	31	Anmerkungen zur Neuauflage des Orientierungsrahmens für den Lernbereich Globale Entwicklung
VIE	34	Neues aus der Kommission/Dialogforum zum Thema SDGs
	36	Rezensionen
	39	Informationen

Sarah Lange

Methodische Reflexionen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien

Zusammenfassung

Der vorliegende Artikel verfolgt das Ziel aufzuzeigen, welche methodischen Herausforderungen mit der Teilnahme von Ländern der Entwicklungszusammenarbeit an Vergleichsstudien verbunden sind. Ausgangspunkt hierfür ist eine detaillierte Analyse und Beschreibung dazu, welche Länder der Entwicklungszusammenarbeit wie häufig und an welchen Zyklen an TIMSS (*Trends in International Mathematics and Science Study*) und PISA (*Programme for International Student Assessment*) teilnahmen. Des Weiteren soll die Rolle von regionalen und nationalen Leistungsmessungen neben den prominenten internationalen Vergleichsstudien (wie TIMSS und PISA) für die Qualitätsentwicklung der Bildungssysteme in Ländern der Entwicklungszusammenarbeit reflektiert werden.

Schlüsselworte: Internationale Leistungsmessung, Entwicklungsländer, Methoden, regionale Vergleichsstudien, PISA, TIMSS

Abstract

This article aims to outline which methodical challenges are related with the participation of development cooperation countries in international large-scale assessment studies. Starting point is a detailed analysis and description which development cooperation countries participated how often in which cycles in TIMSS (*Trends in International Mathematics and Science Study*) and in PISA (*Programme for International Student Assessment*). Furthermore, also the role of regional and national assessment studies in relation to the prominent international large-scale assessments is reflected with view to quality development of educational systems in development cooperation countries.

Keywords: International large-scale assessment, developing countries, methods, regional assessments, PISA, TIMSS

Einleitung

Seit der empirischen Wende in der Erziehungswissenschaft haben die Diskussionen zur Qualitätsentwicklung im Bildungsbereich international zugenommen. Die Auswirkungen dieser Diskussionen lassen sich zunehmend auch in Ländern der Entwicklungszusammenarbeit beobachten. Insbesondere im Kontext der post-2015 Diskussionen, in denen der Fokus stark auf

der Vereinbarkeit von quantitativem (bspw. in Form von mehr Zugang zu Bildungsangeboten) mit qualitativem Fortschritt liegt. Im September 2015 wurden im Rahmen der UN-Resolution „Transforming our world: the 2030 Agenda for Sustainable Development“ auf der UN-Generalversammlung die Sustainable Development Goals verabschiedet (UN General Assembly 2015). Unter diesen 17 Zielen, welche die intensiven internationalen Bestrebungen im Ausbau und in der Entwicklung im Bildungsbereich zwischen 1990 und 2015 im Kontext der *Education for All*-Ziele und der Millennium-Entwicklungsziele (MDGs) zusammenführen und ersetzen (sollen), fokussiert das vierte Teilziel inclusive, gerecht und hochwertige Bildung: „Ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.“ (UNESCO 2015, S. 284). Im Zuge dieser Entwicklungen rückte im letzten Jahrzehnt zunehmend der Output/Outcome eines Bildungssystems – häufig gemessen in Form von Leistungsstudien – als Indikator für Qualität in den Fokus.

Die Entwicklung von gemeinsamen Outcome-Zielen im Bildungsbereich wird stark von der OECD vorangetrieben und kann als Entwicklung einer Form des ‚New Governance‘ beschrieben werden. Mit der Idee von Bildungsmonitoring ist das zentrale Ziel der Systemdiagnostik verbunden (Döbert/Weishaupt 2012, S. 159), d.h. anhand einer kontinuierlich und empirisch gewonnenen Datengrundlage soll „Steuerungshandeln“ begründbarer und zielgerichteter gestaltet werden (ebd.). Damit auf Grundlagen von empirischem Steuerungswissen Qualitätsentwicklung entstehen kann, lassen sich einige Voraussetzungen beschreiben, die erfüllt sein sollten:

- (a) Die gesammelten Daten sind von hoher Qualität und beschreiben ein valides Abbild der sozialen Situation.
- (b) Durch die Auswertung der empirischen Daten lassen sich eindeutige Problembereiche im Bildungssystem identifizieren.
- (c) Die empirischen Ergebnisse werden durch die (lokale) Bildungsforschung für die Öffentlichkeit und die Bildungspolitik verständlich ‚übersetzt‘.
- (d) Die verantwortliche Bildungspolitik nimmt die Ergebnisse wahr und reagiert auf die Problemmarkierung mit angemessenen Maßnahmen (vgl. Kellaghan/Greaney/Murray 2009; Postlethwaite/Kellaghan 2008; Ross/Genevois 2006).

Werden diese Voraussetzungen erfüllt, besteht die Möglichkeit, dass Bildungsmonitoring dazu beitragen kann, dass die Qualität von Bildungsangeboten in einem Bildungssystem verbessert wird. Ob und in wie fern diese Voraussetzungen erfüllt werden, ist maßgeblich von der Konzeption und der Durchführung der Datenerhebung und der Datenanalyse abhängig.

Der vorliegende Artikel verfolgt das Ziel aufzuzeigen, welche methodischen Herausforderungen in Bezug auf die Teilnahme von Ländern der Entwicklungszusammenarbeit an Vergleichsstudien verbunden sind. Ausgangspunkt hierfür ist eine detaillierte Analyse und Beschreibung dazu, welche Länder der Entwicklungszusammenarbeit wie häufig und an welchen Zyklen an TIMSS und PISA teilnahmen. Des Weiteren soll die Rolle der prominenten internationalen Vergleichsstudien (wie TIMSS und PISA) neben regionalen und nationalen Leistungsmessungen für die Qualitätsentwicklung der Bildungssysteme in Ländern der Entwicklungszusammenarbeit reflektiert werden.

Vergleichsstudien mit Ländern der Entwicklungszusammenarbeit

Die Bedeutung von internationalen Vergleichsstudien hat für den Bildungssektor in vielen Ländern seit TIMSS und PISA zugenommen. Die Rezeption der Ergebnisse der Vergleichsstudien variiert stark von Land zu Land. In vielen OECD-Ländern – wie auch in Deutschland – ist die regelmäßige Teilnahme an internationalen Vergleichsstudien fester Bestandteil der qualitätssichernden Maßnahmen im Bildungswesen und Grundlage für eine evidenzbasierte Bildungspolitik geworden. Auffällig ist die zunehmende Teilnahme von Nicht-OECD-Ländern an internationalen Vergleichsstudien. Exemplarisch ist hierfür der Blick auf PISA, der zeigt, dass seit dem ersten PISA-Zyklus an jedem Zyklus mehr sogenannte Partnerländer (Nicht-OECD-Länder) teilnahmen. Im Jahr 2000 waren es vier Länder, bereits 2002 nahmen elf weitere Partnerländer an PISA 2000+ teil (OECD 2013) und zum PISA 2015-Zyklus sind es 38 Länder¹ – und damit mittlerweile mehr teilnehmende Partnerländer als OECD-Länder. Unter den Partnerländern sind auch Länder der Entwicklungszusammenarbeit, die meist durch die finanzielle Unterstützung von Geberinstitutionen der Entwicklungszusammenarbeit (bspw. der Weltbank) teilnehmen (vgl. Liberman/Clarke 2012). Die steigende Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien deutet darauf hin, dass die Bedeutung von Vergleichsstudien und die Rezeption deren Ergebnisse auch in diesem Kontext zunehmen werden.

Methodisches Vorgehen

Im Folgenden wird eine detaillierte Analyse der bisherigen Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien präsentiert. Ziel der Analyse ist es, Muster bezüglich der Teilnahme und der Teilnahmefrequenz dieser Länder nachzuzeichnen. Hierzu wurden in einem ersten Schritt tabellarische Übersichten zu allen Teilnehmerländern an PISA, TIMSS, PIRLS (*Progress in International Reading Literacy Study*), SACMEQ (*The Southern and Eastern Africa Consortium for Monitoring Educational Quality*) und PASEC (*Programme of Analysis of Education Systems*) erstellt. Daraufhin wurden alle Teilnehmerländer an den benannten Studien gruppiert nach den von der Weltbank verwendeten

Einkommensgruppen: Länder mit niedrigem Einkommen (37 Länder), Länder mit niedrigem mittleren Einkommen (47 Länder), Länder mit hohem mittleren Einkommen (58 Länder), Länder mit hohem Einkommen (65 Länder) (UNESCO 2015, S. 317–318)² (vgl. Tabelle 1 & Tabelle 2). Für das Ziel des vorliegenden Aufsatzes, die Teilnahmemuster von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien aufzuzeigen, wurden Tabellen 1 und 2 erstellt, welche Länder der Entwicklungszusammenarbeit zeigen, die im Zeitraum von 1995 bis 2015 an der PISA- und welche an der TIMSS-Studie teilnahmen. Für die folgenden Beschreibungen wurden diese beiden Studien gewählt, da neben dem internationalen Einfluss auch die Frequenz der Durchführungszyklen mit bisher je sechs Zyklen vergleichbar hoch ist. Die Analyseergebnisse zu PIRLS, SACMEQ und PASEC werden als Reflexionsfolien verwendet (vgl. Resümee). Insgesamt zeigt die Übersicht einen deutlichen, wenn auch nicht linearen Anstieg, in der Teilnahme von Ländern der Entwicklungszusammenarbeit an PISA und TIMSS. Bei genauerer Betrachtung zeigen sich deutliche Fluktuationen und Unterschiede dazu, welche Län-

Studie: PISA; Organisation: OECD; Zielgruppe: 15-Jährige; Zyklus: Dreijähriger Zyklus; Tertium Comparationis: Naturwissenschaftliche Kompetenz, Mathematische Kompetenz; Lesekompetenz.		
PISA 2015	LIC:	/
	LMIC:	Georgien, Indonesien, Moldawien, Vietnam.
	UMIC:	Albanien, Algerien, Argentinien, Brasilien, Bulgarien, Kolumbien, Costa Rica, China, Dominikanische Republik, Ungarn, Jordanien, Kasachstan, Libanon, Mazedonien, Malaysia, Mexiko, Montenegro, Peru, Rumänien, Thailand, Tunesien, Türkei.
PISA 2012	LIC:	/
	LMIC:	Indonesien, Vietnam.
	UMIC:	Albanien, Argentinien, Brasilien, Bulgarien, Kolumbien, Costa Rica, Ungarn, Jordanien, Kasachstan, Malaysia, Mexiko, Montenegro, Peru, Rumänien, Serbien, Thailand, Tunesien, Türkei.
PISA 2009	LIC:	Kirgisistan.
	LMIC:	Georgien (2010) ¹ Indien (2010), Indonesien, Moldawien (2010).
	UMIC:	Albanien, Argentinien, Aserbaidshan, Brasilien, Bulgarien, Kolumbien, Costa Rica (2010), Ungarn, Jordanien, Kasachstan, Malaysia (2010), Mauritius (2010), Mexiko, Miranda (Venezuela) (2010), Montenegro, Panama, Peru, Rumänien, Serbien, Thailand, Tunesien, Türkei.
PISA 2006	LIC:	Kirgisistan.
	LMIC:	Indonesien.
	UMIC:	Argentinien, Aserbaidshan, Brasilien, Bulgarien, Kolumbien, Ungarn, Jordanien, Mexiko, Montenegro, Rumänien, Serbien, Thailand, Tunesien, Türkei.
PISA 2003	LIC:	/
	LMIC:	Indonesien.
	UMIC:	Brasilien, Ungarn, Mexiko, Serbien und Montenegro Serbien, Thailand, Tunesien, Türkei.
PISA 2000	LIC:	/
	LMIC:	Indonesien.
	UMIC:	Albanien, Argentinien, Brasilien, Bulgarien, Ungarn, Mazedonien, Mexiko, Peru, Rumänien, Thailand.

Anmerkungen:

1(2010): Nachträglich an der Studie teilgenommen;

LIC = Low income countries, LMIC = Lower middle income countries, UMIC = Upper middle income countries.

Tab. 1: Teilnehmerländer an PISA 2000 bis PISA 2015 mit niedrigem oder mittlerem Einkommensniveau (Quelle: eigene Darstellung)

Studie: TIMSS; Organisation: IEA; Zielgruppe: Vierte und achte Klasse; Zyklus: Vierjähriger Zyklus; Tertium Comparationis: Naturwissenschaftliche Kompetenz, Mathematische Kompetenz.		
TIMSS 2015	LIC:	/
	LMIC:	Armenien, Ägypten, Georgien, Indonesien, Marokko, Palästina.
	UMIC:	Argentinien, Botswana, Bulgarien, Chinesisch Taipeh, Ungarn, Iran, Jordanien, Kasachstan, Libanon, Malaysia, Serbien, Südafrika, Thailand, Türkei.
TIMSS 2011	LIC:	/
	LMIC:	Armenien, Georgien, Ghana, Honduras, Indonesien, Marokko, Palästina, Syrien, Ukraine, Jemen.
	UMIC:	Aserbaidschan, Botswana, Chinesisch Taipeh, Ungarn, Iran, Jordanien, Kasachstan, Libanon, Mazedonien, Malaysia, Rumänien, Serbien, Südafrika, Thailand, Tunesien, Türkei.
TIMSS 2007	LIC:	/
	LMIC:	Armenien, Ägypten, El Salvador, Georgien, Ghana, Indonesien, Mongolien, Marokko, Palästina, Syrien, Ukraine, Jemen.
	UMIC:	Algerien, Bosnien und Herzegowina, Botswana, Bulgarien, Chinesisch Taipeh, Kolumbien, Ungarn, Iran, Jordanien, Kasachstan, Libanon, Malaysia, Rumänien, Serbien, Thailand, Tunesien, Türkei.
TIMSS 2003	LIC:	/
	LMIC:	Armenien, Ägypten, Ghana, Indonesien, Moldawien, Marokko, Palästina, Philippinen, Syrien, Jemen.
	UMIC:	Argentinien, Botswana, Bulgarien, Chinesisch Taipeh, Ungarn, Iran, Jordanien, Libanon, Mazedonien, Malaysia, Rumänien, Serbien, Südafrika, Tunesien.
TIMSS 1999	LIC:	/
	LMIC:	Indonesien, Moldawien, Marokko, Philippinen.
	UMIC:	Bulgarien, Chinesisch Taipeh, Ungarn, Iran, Jordanien, Mazedonien, Malaysia, Rumänien, Südafrika, Thailand, Tunesien, Türkei.
TIMSS 1995	LIC:	/
	LMIC:	Indonesien, Philippinen.
	UMIC:	Argentinien, Bulgarien, Kolumbien, Ungarn, Iran, Mexiko, Rumänien, Südafrika, Thailand.

Anmerkungen:

LIC = Low income countries, LMIC = Lower middle income countries, UMIC = Upper middle income countries.

Tabelle 2: Teilnehmerländer an TIMSS 1995 bis TIMSS 2015 mit niedrigem oder mittlerem Einkommensniveau.

der und welche Regionen an welcher der beiden Studien einzeln, mehrfach oder regelmäßig teilnahmen. Ausgehend von dieser Übersicht wurden Analysen anhand von zwei Analyseeinheiten durchgeführt. Zum einen wurden die Teilnehmerländer analysiert nach geografischer Weltregion (Analyseeinheit Region) und zum anderen mit Blick auf die Teilnahmehäufigkeit dieser Länder (Analyseeinheit Teilnehmerland).

Analyse der Teilnehmerländer nach Regionen (1995–2015)

Die Analyse der PISA- und TIMSS-Teilnehmerländer mit niedrigem und mittlerem Einkommen an den Studien wurde betreffend ihrer geografischen Region anhand der von der UNESCO vorgeschlagenen EFA Regionen³ durchgeführt (UNESCO 2014, S. 316–217).

Subsahara-Afrika: In dieser Region nahmen bisher vier der 46 Länder an TIMSS und oder PISA teil. Auffallend ist, dass aus

Subsahara Afrika bisher nur Mauritius einmal (2009) an PISA teilgenommen hat. Dies ändert sich erst mit dem PISA for Development-Projekt, an dem Sambia und Senegal teilnehmen (vgl. Ward in diesem Heft). Dagegen haben an TIMSS bisher drei subsaharische Länder mehrfach teilgenommen (Südafrika an fünf Zyklen, Ghana an vier Zyklen und Botswana an zwei Zyklen).

Arabische Staaten: In dieser Region nahmen bisher neun der 20 Länder an TIMSS und/oder PISA teil. Die Analyse der arabischen Staaten zeigt einen auffallenden regionalen Trend dazu an TIMSS und weniger an PISA teilzunehmen. Von den neun arabischen Staaten, die mit Ausnahme von Algerien alle mindestens an drei TIMSS-Zyklen teilgenommen haben, haben nur zwei Länder (Jordanien und Tunesien) auch mehrfach an PISA teilgenommen.

Zentral- und Osteuropa: In dieser Region nahmen bisher zehn der 21 Länder an TIMSS und/oder PISA teil. Die zentral- und osteuropäischen Länder nehmen intensiv an beiden Studien teil. Auffallend ist, dass von den zehn Teilnehmerländern sieben Länder an TIMSS und PISA teilnehmen und dass die Konstanz der Teilnahme sehr hoch ist. Von den Teilnehmerländern dieser Gruppe hat jedes Land durchschnittlich an 3,5 Zyklen von PISA und/oder TIMSS teilgenommen.

Zentralasien: In dieser Region nahmen bisher sechs der neun Länder an TIMSS und/oder PISA teil. Insgesamt scheint in dieser Region vergleichbar spät der Trend zur Teilnahme an Vergleichsstudien aufgenommen worden zu sein, da (außer Armenien) alle Länder erst ab dem dritten oder vierten Zyklus teilnahmen – und dann relativ ausgewogen an PISA und/oder TIMSS. Die durchschnittliche Teilnahme der Länder dieser Gruppe ist mit zwei Zyklen vergleichsweise niedrig.

Ostasien und Pazifik: In dieser Region nahmen bisher elf der 33 Länder an TIMSS und/oder PISA teil. In dieser Region sind es sieben Länder, die an PISA und drei Länder, die an TIMSS partizipierten. Die Konstanz der teilnehmenden Länder ist sehr hoch – im Durchschnitt haben die Teilnehmerländer dieser Region an 4,2 Zyklen partizipiert. China hat im PISA 2015 Zyklus das erste Mal an einer Vergleichsstudie teilgenommen. Indonesien, Malaysia und Thailand weisen seit Beginn regelmäßige Teilnahmen an PISA und TIMSS auf. Die Präferenz dieser Region tendiert im Gesamt zur Teilnahme an der PISA-Studie.

Lateinamerika und die Karibik: In dieser Region nahmen bisher elf der 43 an TIMSS und/oder PISA teil. Die Analyse der Teilnehmerländer aus Lateinamerika und in der Karibik zeigt, dass auch diese Region deutliche Tendenz zur Teilnahme an PISA zeigt. Insbesondere Brasilien und Mexiko stechen dabei heraus, da beide Länder an allen bisherigen sechs PISA-Zyklen teilgenommen haben. Nur Argentinien nahm bisher regelmäßig an PISA und an TIMSS statt. Auffällig ist, dass es einige Länder gibt, die aus dieser Region nur vereinzelt an den Studien teilnahmen – bspw. Honduras 2011 an TIMSS oder Panama 2009 an PISA.

Süd- und Westasien: Von den neun Ländern, die zu Südwestasien zählen, haben bisher zwei Länder an Vergleichsstudien teilgenommen: zum einen der Iran mit einer regelmäßigen Teilnahme an allen TIMSS-Zyklen und zum anderen Indien mit einer bisher einmaligen Teilnahme in 2009 an PISA.

Fazit: Insgesamt zeigen sich deutliche regionale Unterschiede. Der Überblick (Abb. 1) ermöglicht es regionale Präferenzen

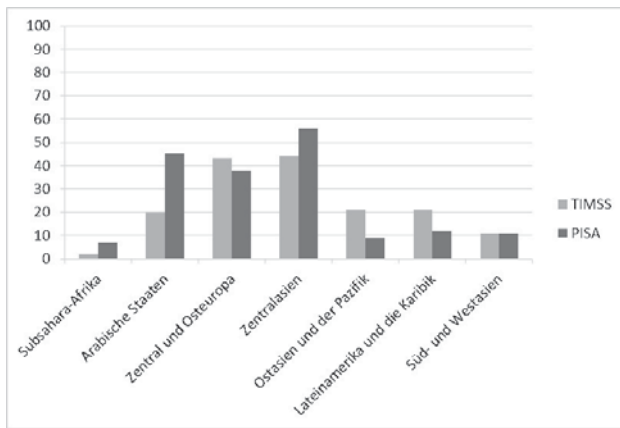


Abb. 1: Anteil (in %) der teilnehmenden Länder mit niedrigem und mittlerem Einkommen nach Region (Quelle: Eigene Darstellung)

dazu abzulesen, ob Länder in einer Region häufiger an PISA oder an TIMSS teilnehmen. Zudem zeigt die Abbildung wie groß der Anteil der teilnehmenden Länder mit niedrigem und mittlerem Einkommen ist – relativiert an der Anzahl der Länder pro Region. Diese Präferenzen können unterschiedlich interpretiert werden. Denkbar ist, dass dies der Ausdruck von überregionalen Assessment-Strategien ist, die auch Hypothesen zu dem erwarteten Nutzen widerspiegeln. Zudem kann der Wunsch sich mit benachbarten Ländern zu vergleichen ein denkbarer Anreiz für die Teilnahme an derselben Studie sein. So nehmen Länder in Subsahara-Afrika und arabische Länder mit niedrigen und mittleren Einkommen eher an TIMSS teil – dagegen nehmen die teilnehmenden Länder aus Ostasien sowie aus Lateinamerika eher an den PISA-Zyklen teil.

Analyse der Teilnehmerländer nach Teilnahmehäufigkeit (1995–2015)

Im Folgenden wird die Frage verfolgt, ob sich neben den regional andeutenden Strategien auch auf Länderebene nationale Assessment-Strategien betreffend der Nutzung von Vergleichsstudienten zeigen. Dazu werden die Länder hinsichtlich ihrer Teilnehmerhäufigkeit betrachtet – insbesondere Länder mit einer hohen Teilnehmerhäufigkeit, Länder mit einer niedrigen Teilnehmerhäufigkeit und Länder, die an beiden Studien teilnehmen.

Hohe Teilnehmerhäufigkeit: An allen sechs PISA-Zyklen nahmen bisher vier Länder teil: Ungarn, Indonesien, Brasilien, Mexiko. An allen sechs TIMSS-Zyklen nahmen bisher teil: Ungarn, Indonesien, Iran.

Einmalige Teilnahme: Nur ein einziges Mal nahmen an PISA bisher die folgenden acht Länder teil: Mauritius (2009), Libanon (2015), Algerien (2015), China (2015), Dominikanische Republik (2015), Panama (2009), Miranda (Venezuela) (2010), Indien (2009). An TIMSS nahmen bisher nur ein einziges Mal die folgenden sechs Länder teil: Algerien (2007), Azerbaijan (2011), El Salvador (2007), Honduras (2011), Mexiko (1995).

Eine nur einmalige Teilnahme kommt in allen Weltregionen vor und lässt verschiedene Interpretationen zu. Es ist möglich, dass die häufigere Teilnahme auf Grund von Finanzierungsschwierigkeiten nicht möglich war. Auch möglich ist, dass eventuell die Rezeption und damit die Nutzung der em-

pirischen Ergebnisse nicht umfassend geschah und dazu führte, dass der Aufwand einer Teilnahme aus Sicht der nationalen Regierung als nicht lohnend angesehen wurde. Ein weiterer Grund für die nicht-Kontinuität kann sein, dass die Teilnahme an einer internationalen Studie federführend von Gebern initiiert und angetrieben wurde und weniger oder gar nicht von Seiten des teilnehmenden Landes (Jansen 2005, S. 21; vgl. den Fall Kirgistan: OECD 2013, S. 29).

Intensive Teilnahme an PISA und TIMSS: Insgesamt haben die folgenden neun Länder an vier oder mehr Zyklen von beiden Studien teilgenommen: Jordanien, Tunesien, Bulgarien, Ungarn, Rumänien, Serbien, Türkei, Indonesien, Thailand.

Fazit: Die Teilnahme an beiden großen Studien weist darauf hin, dass in diesen neun Ländern mit niedrigen und mittleren Einkommen den Ergebnissen von Vergleichsstudien eine hohe Bedeutung beigemessen wird. Gründe für die insgesamt deutlich beobachtbare Zunahme der Teilnahme und der Teilnahmehäufigkeit an Leistungsstudien gehen sicher auch mit der gestiegenen Bedeutung der Rechenschaftsablegung (accountability) und der Evidenzbasierung im Bildungsbereich einher (vgl. Benavot/Köseleci 2015, S. 3).

Resümee

Die Analyseergebnisse zur Teilnahme an internationalen Vergleichsstudien werden im Folgenden in Verbindung gesetzt mit regionalen und nationalen Vergleichsstudien, die in Ländern der Entwicklungszusammenarbeit durchgeführt werden. Die Ergebnisse zeigen eindeutige Unterschiede in der Teilnahmeintensität im Vergleich der Einkommensgruppen der Teilnehmerländer. So ist Kirgistan, das 2006 und 2009 an PISA teilgenommen hat, das bisher einzige Teilnehmerland aus der Gruppe der Länder mit niedrigem Einkommen. An der regionalen Vergleichsstudie SACMEQ dagegen haben aus dieser Gruppe bisher sieben Länder teilgenommen – darunter bspw. Kenia, Malawi oder Zimbabwe.

Im Vergleich der Gruppen der Länder mit mittlerem Einkommen sind es auffallend wenig Teilnehmerländer, die aus der Untergruppe der Länder mit niedrigem mittleren Einkommen an internationalen Vergleichsstudien teilnehmen (PISA: z.B. Indonesien, Moldavien, Indien; TIMSS: z.B. Armenien, Ghana, Ägypten). Dies legt nahe, dass die Zugangshürden für Länder mit niedrigem und mit niedrigem mittleren Einkommen für die Teilnahme an internationalen Vergleichsstudien höher sind als für die Teilnahme an regionalen Vergleichsstudien. In dem Evaluationsbericht des IEA-Programms PASA (*Program on the Assessment of Student Achievement*) empfiehlt Lockheed (2010, S. vii) „IEA concentration on building capacity among lower-middle income countries rather than intensifying efforts in low-income countries. The technical requirement for participating successfully in IEA studies is often beyond the existing capacity in low income countries, while lower-middle income countries frequently have a suitable basis on which to build.“

Die deskriptiven Analysen zeigen, dass betreffend der Teilnahme an internationalen Vergleichsstudien die Region Subsahara-Afrika verglichen mit anderen Weltregionen in der Teilnahmeintensität und -häufigkeit deutlich zurück liegt. Interessant ist, dass Länder dieser Region, die bisher an PISA oder TIMSS teilgenommen haben, auch intensiv an regionalen Ver-

gleichsstudien teilnehmen. So haben Südafrika, Botswana und Mauritius je an drei der vier Zyklen von SACMEQ teilgenommen. Dies lässt vermuten, dass die Teilnahme an regionalen Vergleichsstudien Interesse sowie personelle und technische Kapazitäten für die Teilnahme an internationalen Vergleichsstudien erhöht. Auch im Hinblick auf die *PISA for Development*-Länder scheint dieser Zusammenhang naheliegend. So haben Ecuador, Guatemala und Paraguay je an SERCE (*Segundo Estudio Regional Comparativo y Explicativo*) und an TERCE (*Tercer Estudio Regional Comparativo y Explicativo*) – dem zweiten und dritten Zyklus von LLECE⁴ – teilgenommen. Zambia nahm an allen vier SACMEQ-Zyklen teil und Senegal hat an PASEC 2006 und 2014 sowie an MLA I und II teilgenommen. Auch Indonesien, Ungarn und der Iran – Länder, die durch ihre hohe Teilnehmerfrequenz an TIMSS oder PISA auffallen – nehmen parallel intensiv an PIRLS teil. Ausgehend von den Ergebnissen der Analysen lässt sich die Hypothese aufstellen, dass die Teilnahmen an nationalen, regionalen und internationalen Vergleichsstudien zusammenhängen und sich gegenseitig bedingen.

Methodische Reflexionen

Herausforderungen methodischer Art lassen sich bei Vergleichsstudien viele benennen. Diese lassen sich systematisieren hinsichtlich der Durchführung und der Rezeption der Ergebnisse eines Forschungsprojekts (vgl. im Überblick: Ross/Genevois 2006). Im Zuge der zunehmenden Diskussionen zur Bedeutung von Vergleichsstudien haben sich viele Autorinnen und Autoren diesen allgemeinen Herausforderungen gewidmet, bspw. die Herausforderung von Messäquivalenz bei *cross-cultural studies* (vgl. Mullen 1995) oder der damit verwandten Konzepten der Kultur- und Testfairness (vgl. Anastasi 1964) sowie allgemeiner zu Methoden des (Kultur-)Vergleichs (vgl. Rademacher 2013).

Bezugnehmend auf die Analysen zur Teilnahme von Ländern der Entwicklungszusammenarbeit an den internationalen Vergleichsstudien PISA und TIMSS, werden im Folgenden zwei Aspekte fokussiert, die spezifisch sind für methodische Herausforderungen mit Blick auf Länder der Entwicklungszusammenarbeit. Hierzu gehören die Leistungsmessung auf den unteren Kompetenzstufen und zielgruppenadäquate Kommunikationsstrategien der Ergebnisse.

Leistungsmessung auf den unteren Kompetenzstufen

Für die Messung von kognitiven Fähigkeiten sind – wie für jede Form der Messung – zuverlässige Daten von zentraler Bedeutung. Ausschlaggebend ist hierfür insbesondere die Testkonstruktion. Mit Blick auf Länder der Entwicklungszusammenarbeit ist insbesondere die Konstruktion eines Leistungstests, der auch auf den unteren Kompetenzstufen valide misst, eine Herausforderung. Die vorliegenden Ergebnisse von kognitiven Leistungstests zeigen, dass in Ländern der Entwicklungszusammenarbeit der Anteil der Schülerinnen und Schüler auf niedrigen Leistungsniveaus sehr groß ist. Hinweise darauf zeigen überregionale Vergleichsstudien wie SACMEQ, PASEC oder MLA (*Monitoring Learning Achievement*) in Ländern in Subsahara Afrika oder wie LLECE in Südamerika. In sieben Ländern in Subsahara-Afrika, die an SACMEQ II teilgenommen

haben, zeigten die Ergebnisse, dass mehr als 40 % der teilnehmenden Schülerinnen und Schüler nicht lesen und nicht schreiben können (vgl. Lee/Zuze/Ross 2005; Smith/Barrett 2011, S. 24). Auch die Ergebnisse von nationalen Assessment-Studien weisen auf niedrige Leistungsergebnisse und auf die Notwendigkeit, Wege zu suchen das ‚Können‘ und ‚Wissen‘ von Schülerinnen und Schülern zu verbessern (Benavot/Köseleci 2015, S. 10).

Auch bisherige PISA-Studien zeigen, dass Partnerländer – also nicht OECD-Länder – verglichen mit dem OECD-Durchschnitt deutlich niedrigere Leistungsergebnisse aufweisen. Im OECD Education Working Paper No. 93 werden in einer Grafik die Leistungsergebnisse und der Anteil der Schülerinnen und Schüler, die die Kompetenzstufe 2 im Lesen nicht erreichen, im Vergleich aller Partnerländer (inklusive der Türkei, Chile und Mexiko – den OECD-Ländern mit mittleren Einkommen) in einer Grafik zusammen mit dem jeweiligen OECD Durchschnitt dargestellt. Das Fazit des Vergleichs lautet: „In the majority of partner countries, the share of students below Level 2 proficiency in reading exceeds 40%, compared to the OECD average of 19%.“ (OECD 2013, S. 13). Mit Blick auf diese Vergleiche stellt sich die Frage, welche Bedeutung der Vergleich der Leistungsergebnisse von Ländern der Entwicklungszusammenarbeit gemessen am Maßstab des OECD-Durchschnitts hat. Wie fair ist es, an dem OECD-Maßstab Länder zu messen, deren durchschnittlichen Leistungsergebnisse nicht in die Berechnung des OECD-Mittelwerts mit einfließen? Sollten die *PISA for Development*-Länder nicht an einem entsprechend separaten Mittelwert verglichen werden?

Die für bspw. die PISA-Studie entwickelten Instrumente messen die Leistung derjenigen Schülerinnen und Schüler am genauesten, die der Population mit dem durchschnittlichen Leistungsniveau angehören. Wenn der überwiegende Teil der untersuchten Zielgruppe eines Landes jedoch auf den unteren Kompetenzstufen verweilt, kann dies die Qualität der Messungen der PISA-Tests für dieses Land beeinträchtigen und zu Messungenauigkeit führen (OECD 2013, S. 15).

Im Rahmen des Projekts der OECD, PISA in Entwicklungsländern durchzuführen, soll diese Schwäche bisher entwickelter Erhebungsinstrumente verbessert werden, indem neue Items ergänzt werden und dabei die kognitiven Testinstrumente anhand der PISA-Skala vergleichbar bleiben (vgl. Beitrag von Ward in diesem Heft; Adams/Cresswell 2016). Eine zentrale methodische Herausforderung wird es für *PISA for Development* dabei sein, die Balance zu wahren zwischen der Veränderung und der Anpassung von PISA-Instrumenten für den Einsatz in Ländern der Entwicklungszusammenarbeit und der gleichzeitigen Weiterentwicklung von internationalen Bildungsindikatoren.

Ein Schritt in diese Richtung war die für PISA 2012 eingeführte „Easy Booklet Option“. Jedes an PISA teilnehmende Land kann diese Option wählen, um anhand zusätzlicher Items auf den unteren Kompetenzlevels dazu beizutragen, dass die Messung von erwarteten niedrigeren Leistungsergebnissen zuverlässiger verläuft. Jedoch wurde diese Option in PISA 2009 und PISA 2009+ nur von 17 Ländern gewählt – möglicherweise auf Grund von fehlendem Bewusstsein für den Stand der nationalen Leistungen oder auf Grund von Angst vor Stigmatisierung (vgl. OECD 2013, S. 15). Als

Pendant zur ‚Easy Booklet‘-Option bei PISA verfolgt die IEA folgende Strategie für interessierte neue Mitgliedsländer. Länder der Entwicklungszusammenarbeit, die an PIRLS teilnehmen möchten, haben die Möglichkeit an pre-PIRLS zu partizipieren – eine Studie mit derselben Konzeption wie PIRLS, die jedoch weniger schwierige Testitems aufweist. Explizites Ziel von pre-PIRLS ist es, die Bildungspolitik in Ländern der Entwicklungszusammenarbeit dazu zu informieren, welche Schritte unternommen werden sollten, um die Teilnahme an PIRLS zu ermöglichen. Die regionale Studie SACMEQ ermöglicht es während des aktuellen Zyklus (SACMEQ IV) Angola in der Beobachterrolle teilzunehmen, mit der Aussicht ein Vollmitglied zu werden.⁵

Zielgruppenadäquate Kommunikationsstrategien der Ergebnisse

Eine weitere Herausforderung insbesondere für an Vergleichsstudien teilnehmende Länder der Entwicklungszusammenarbeit ist die zielgruppenadäquate Kommunikation der empirischen Ergebnisse einer Studie. Mit Blick auf die sehr unterschiedlichen Zielgruppen, für die die Ergebnisse von Vergleichsstudien wissenswert sein können sowie betreffend deren Interessen und Anliegen, bedarf es der spezifischen Aufbereitung der Ergebnisse. Relevante Zielgruppen sind hierbei insbesondere die nationale Bildungspolitik, im Bildungsbereich Tätige (bspw. Lehrkräfte, Schulleitende), die interessierte Öffentlichkeit (bspw. Eltern) sowie die lokale scientific community. Es bedarf gezielter Kommunikationsstrategien, um sicherzustellen, dass die Studienergebnisse korrekt bspw. in den Medien wiedergegeben werden. Hierzu kann idealerweise die lokale Bildungsforschung beitragen, indem neben den Publikationen der internationalen Vergleiche Bildungsforschende gezielt den Blick auf die vor der Studie formulierten Erkenntnisinteressen der lokalen Bildungspolitik wirft. Dies kann zum einen die Aufbereitung der Ergebnisse und der gegebenenfalls sinnvollen Übersetzung in lokale Sprachen bedeuten. Zum anderen sind es erst weiterführende Sekundäranalysen, die Antworten liefern können auf spezifischere Fragen zur Situation des nationalen Bildungssystems. Die Erfahrungen mit PISA in den letzten fünfzehn Jahren haben gezeigt, dass mit der Teilnahme von Partnerländern klar das Ziel verbunden und verfolgt werden sollte, institutionelle Kapazität für Assessment und für Auswertungen aufzubauen (vgl. Lockheed 2013).

Wenn die wissenschaftliche Begleitung und die Durchführung von Sekundäranalysen in Länder der Entwicklungszusammenarbeit häufig gar nicht oder nur sehr begrenzt stattgefunden haben, dann lag das in den meisten Fällen an den mangelnden technischen und personellen Ressourcen der lokalen Bildungsforschung (vgl. OECD 2013, S. 20–27). Adair (1995, S. 646) schlägt die folgenden vier Kategorien zur Gruppierung von Determinanten vor, die die Umgebung von universitärer Forschung (im Fachbereich Psychologie) beeinflussen: (1) materiale und personelle Ressourcen, (2) professionelle Interaktionen und Kommunikation, (3) Forschungsrelevanz und Auszeichnungen, (4) akademische Führung. Das nur begrenzte Vorhandensein von Ressourcen hat aus seiner Sicht den offensichtlichsten und direkten Effekt auf (ausbleibende) Forschungsaktivitäten (Adair 1995, S. 644; vgl. Wedgewood/Hammett 2005, S. 7).

Die Planung für die zielgruppenspezifische Kommunikation der empirischen Ergebnisse beginnt bereits in der Konzeptionsphase einer Studie. Laut Postlethwaite (2006, S. 107–108) ist ein Qualitätskriterium für qualitativ hochwertige Vergleichsstudien, dass zu Beginn einer Studie das Erkenntnisinteresse von Seiten der lokalen Bildungspolitik eingeholt wird; denn wenn Daten erhoben werden, die auf diese Belange antworten erhöht dies die Wahrscheinlichkeit, dass der Bericht gelesen und wahrgenommen wird.

Murimba (2005) berichtet davon, welche Effekte die SACMEQ-Studie auf die Teilnehmerländer hat. In den vier Ländern (Botswana, Lesotho, Seychellen, Malawi), in denen die Durchführung von SACMEQ I zum personellen Kapazitätsaufbau innerhalb des Landes geführt hat (bspw. in Form von intensiven Weiterbildungen) hat die Datenaufbereitung für SACMEQ II mit vier bis zehn Monaten deutlich kürzer gedauert als in Ländern, die zu SACMEQ II das erste Mal an der Studie teilnahmen (wie bspw. Mosambik, das 24 Monate für die Datenaufbereitung benötigte). Des Weiteren wurde die technische Ausstattung (bspw. leistungsstarke Computer) in z.B. Tansania auch für die Durchführung von Forschungsprojekten im Bildungsbereich genutzt oder für die Einführung von zuverlässigen nationalen Assessment-Studien. Auch bewirkte die Teilnahme an SACMEQ in vielen Ländern eine deutliche Stärkung der nationalen Bildungspolitik (vgl. Murimba 2005, S. 94–98). Lockheed (2015, S. vii) beschreibt positive Effekte der Teilnahme an internationalen Vergleichsstudien, insbesondere betreffend Curriculumsreformen.

Kommunikation ist nicht nur bezüglich der Verbreitung der Ergebnisse relevant, sondern auch vor und während der Datenerhebung. Es ist eine ethische Herausforderung für die Durchführung von empirischer Forschung, wenn die ‚Beforschten‘ eine Personengruppe sind, bei denen nicht davon ausgegangen werden kann, dass grundlegende Vorstellungen zu Forschung, Forschungszielen, der Funktion von Datenerhebungsinstrumenten sowie zu der Art und Weise wie gewonnenen Daten weiterverwendet werden vorhanden sind. Voraussetzung für Forschung im Kontext von Ländern der Entwicklungszusammenarbeit ist daher größtmögliche Transparenz und dezidierte Einführungen in die Abläufe eines Forschungsprozesses und die Bedeutung und weitere Verwendung von empirischen Daten (vgl. weiterführend: Hammett/Wedgewood 2005; Lang-Wojtasik 2002).

Die korrekte Rezeption von Vergleichsstudienergebnissen kann auch in OECD-Ländern eine Herausforderung sein. Der Umgang mit dem PISA-Schock in 2000 in Deutschland zeigt, dass es durchaus eine längere ‚Lerngeschichte‘ benötigen kann bis eine Gesellschaft einen differenzierteren Umgang mit Ergebnissen aus Vergleichsstudien aufweist und nicht das scheinbare Argument „Wie PISA zeigt...“ für so ziemlich jede bildungsbezogene Aussage verwendet wird. Am Fall der Rezeption der PISA-Ergebnisse in Deutschland verweisen Prenzel/Drechsel (2004) auf die hohe Bedeutung einer differenzierten Rezeption und auf die Notwendigkeit, die Reichweite der Aussagen von Vergleichsstudien klar im Blick zu behalten (vgl. Schwippert/Bos 2005).

Diesem Problem wurde im Rahmen von bspw. SACMEQ entgegengewirkt, indem der lokale Kapazitätsauf- und -ausbau explizit gesetzte und angestrebte Ziele der überregionalen Ver-

gleichsstudie sind und dies auch zu Erfolgen führte (vgl. Murimba 2005). Im Rahmen von PISA for Development wird dieses Problem in Angriff genommen, indem als Teil der Konzeptionsphase in jedem teilnehmenden Land analysiert wird, welche Kapazitäten vorhanden sind, welche noch benötigt werden und wie und in welchem zeitlichen Rahmen diese aufgebaut werden (vgl. dazu den Bericht zur Kapazitätenanalyse von Sambia: OECD/World Bank Group 2014).

National versus international: Wem nutzt welche Studie?

Für die Frage des Nutzens von Vergleichsstudien ist es nicht trivial die verschiedenen Arten von Vergleichsstudien zu unterscheiden. Zur Systematisierung ist die Frage hilfreich, welche Regionen oder Länder anhand des festgelegten Tertium Comparationis der jeweiligen Studien miteinander verglichen werden; diese Frage liegt der folgenden Unterscheidung zugrunde: *a) Nationaler Vergleich:* Der Vergleich verschiedener Teilpopulationen innerhalb eines Landes miteinander (z.B. nationale Assessments wie die Vergleichsarbeiten (VERA) in Deutschland (vgl. Maier/Kuper 2012) oder die National Annual Assessment (ANA) in Südafrika⁶ (vgl. Benavot/Köseleci 2015, S. 18).

b) Regionaler Vergleich: Der Vergleich ausgewählter Länder einer Region, die gewisse Gemeinsamkeiten aufweisen (z.B. SACMEQ: seit 1995 Vergleich der Kompetenzen von Schülerinnen und Schülern in Ländern im südlichen und östlichen Afrika, die alle Englisch als Amtssprache haben⁷).

c) Internationaler Vergleich: Der großangelegte Vergleich von Ländern aus unterschiedlichen Weltregionen – vornehmlich organisiert von Seiten der OECD oder der IEA. Die anfängliche Gemeinsamkeit der OECD-Zugehörigkeit der Länder, die an PISA teilnahmen, wurde zunehmend aufgebrochen und mehr und mehr Partnerländer mit aufgenommen.

Zu der Frage wer ‚wem nutzt welche Studie‘ bezieht Benavot als Leiter des Global Monitoring Reports klar Position zugunsten nationaler Vergleiche. Mit sogenannten *national learning assessments* sind nationale Vergleichsstudien gemeint, anhand der Lernergebnisse mit gesetzten nationalen Zielen der Bildungsbehörden überprüft werden. Im Unterschied zu *high stakes-Tests* werden die Ergebnisse nicht für bspw. die Selektion an Übergängen im Bildungswesen verwendet (vgl. Benavot/Köseleci 2015, S. 3–4). Als Vorteil für die Teilnahme an nationalen Assessment-Studien nennen die Autoren die größere Kontextsensibilität neben der wie bei internationalen Vergleichsstudien gegebenen Möglichkeit auf nationaler Ebene die Daten für Bildungsmonitoring zu verwenden (vgl. Benavot/Köseleci 2015, S. 18).

Abhängig vom Entwicklungsstand eines Bildungssystems mag die Teilnahme an einer nationalen Assessment-Studie bevorzugt werden, da fachorientiert und basierend auf dem nationalen Curriculum überprüft wird, ob die Schülerinnen und Schüler die Inhalte, die gelehrt werden (sollen) auch können. Zudem sind nationale Studien flexibel mit Blick auf die untersuchte Zielgruppe. Die Zielgruppe der 15-Jährigen mag für manche Länder nicht im Hauptfokus liegen. Interessant hierzu ist, dass die meisten nationalen Assessments zwischen 1990 und 2013 in den Klassenstufen 4–6 durchgeführt wurden (vgl. Benavot/Köseleci 2015, S. 9). Benavot/Köseleci (2015) vertreten die Position, dass es nicht den internationalen Ver-

gleich bedarf, sondern lediglich die Durchführung und die Bereitstellung empirischer Daten, um daraus Steuerungswissen für die lokal verantwortlichen Entscheidungsträger generieren zu können, um zur Qualitätsentwicklung des Landes beizutragen.

Kellaghan/Greaney/Murray (2009) beschreiben detailliert die wichtigsten Schritte, auf die geachtet werden sollte, um eine nationale Assessment-Studie für die Qualitätsentwicklung zu verwenden. Nationale Leistungsmessungen bieten mehr noch die Möglichkeit Datenerhebungsinstrumente mit den spezifischen Interessen der lokalen Bildungspolitik abzustimmen. Von Seiten der UNESCO wird die Durchführung von nationalen Assessment-Studien als Indikator gesehen, um die internationalen Bildungsziele voranzubringen und vor allem für die nationale Qualitätsentwicklung (vgl. Übersicht der nationalen Assessment-Studien zu jedem Land der Welt: UNESCO 2015, S. 304–311). Ein anderes Argument ist das der Sprache: nationale Assessments können leichter in den bedeutendsten Sprachen in einem Land durchgeführt werden. So führt Südafrika aktuell eine nationale Assessment-Studie in zwölf verschiedenen Sprachen durch (vgl. Benavot/Köseleci 2015, S. 18; vgl. Naumann 2005, S. 235).

Neben diesen Argumenten für nationale Assessment-Studien stellt sich jedoch die Frage nach der Qualität solcher Studien und nach den technischen Kapazitäten mit denen die Studien durchgeführt wurden. Benavot/Köseleci (2015, S. 4) verweisen darauf, dass die methodische Qualität und die wissenschaftlichen Standards der Studien bislang noch nicht analysiert werden konnten. Für die Teilnahme an regionalen oder internationalen Vergleichsstudien spricht das Argument, dass damit Zugang zu intensiver Schulung und zu einem hohen Maß an technischem Knowhow und damit verbundenen Chancen zum lokalen Kapazitätsausbau für die Durchführung und Analysen von empirischen Daten verbunden sind. Die im Rahmen der Teilnahme an großen Vergleichsstudien gewonnenen Kompetenzen sowie technischer Ausstattungen können weiterführend für den Ausbau der lokalen empirischen Bildungsforschung genutzt werden.

Zugunsten von regionalen und internationalen Vergleichen spricht das Ziel, universelle Bildungsindikatoren zu formulieren. „The world now faces a challenge to define global learning indicators that can be measured and tracked on a global scale over time.“ (Ward/Zoido in diesem Heft, S. 21–25) Für dieses Ziel bedarf es der Entwicklung einer gemeinsamen Skala, auf der die Ergebnisse weltweit erhobener Schülerleistung abgetragen werden können. Aufbauend auf bisherigen Erfahrungen mit regionalen und internationalen Leistungsmessungen verfolgt das UNESCO Institut für Statistik das Ziel Kriterien zu entwickeln, anhand derer der Stand des Lernerfolgs betreffend Kompetenzen im Bereich Lesen und im Rechnen gemessen werden kann. „This is an excellent opportunity for the learning assessment and educational development communities to work together. Effective collaboration and sharing of knowledge at this stage will lay the foundation for a system that supports all countries in their efforts to measure their progress towards achieving a post-2015 learning goal.“⁸

Ausblick

Als Ausblick zu diesem umfassenden Thema werden im Folgenden vier Gedanken thematisiert, die in weiterführenden

Debatten zur Teilnahme von Ländern der Entwicklungszusammenarbeit an Vergleichsstudien berücksichtigt werden sollten: (a) *Weiterer Forschungsbedarf zum Nutzen von Vergleichsstudien*: Zu der Frage nach dem Nutzen der Teilnahme an Vergleichsstudien für Länder der Entwicklungszusammenarbeit gibt es noch großen Forschungsbedarf. Für eine weiterführende Diskussion bedarf es mehr Evidenzbasierung, bspw. in Form von Evaluationsstudien, die ausgewählte Länder, die intensiv an nationalen, regionalen und/oder internationalen Vergleichsstudien teilgenommen haben, in den Blick nehmen: Wie war die Rezeption der Vergleichsstudien? Welche bildungspolitischen Reformen wurden auf Grund der Ergebnisse eingeführt und welche Effekte haben diese auf das Bildungssystem des Landes? (b) *PISA versus PISA for Development*: Des Weiteren wird es zu beobachten bleiben, welche Rolle *PISA for Development* für die Qualitätsentwicklung der Teilnehmerländer spielen wird – insbesondere im Vergleich zu Ländern der Entwicklungszusammenarbeit, die sich (weiterhin) für eine Teilnahme an den regulären PISA oder TIMSS-Zyklen entscheiden. Erst weitere Analysen zur Prüfung der Qualität von nationalen Assessment-Studien noch mehr Auskunft zum Gehalt dieser sich ausweitenden Alternativform des Generierens von empirischen Daten für eine evidenzbasierte Bildungspolitik liefern.

(c) *Zielgerichteter Austausch von Expertwissen*: Mehr Zusammenarbeit und gegenseitiger Austausch zwischen den Akteuren zu dem jeweils gewonnenen methodischen Expertenwissen kann zur Entwicklung der technischen Qualität der Vergleichsstudien in Ländern der Entwicklungszusammenarbeit beitragen. Interessant für Austausch sind gerade auch Erfahrungen im Kontext von formativen Messungen durch Haushaltserhebungen wie ASER (Annual Status of Education Report) in Indien oder auch EGMA (Early Grade Mathematics Assessment) und EGRA (Early Grade Reading Assessment), deren direkte Analysen vor Ort das Ziel verfolgen mit passgenauen Förderprogrammen dort anzusetzen, wo Bedarf vorherrscht. (d) *Die Chancen von Open Data-Strategien für Forschung*: Die Datenfreigabestrategien von bspw. SACMEQ und von der OECD für *PISA for Development* bieten für die scientific community im Bereich Bildung und Entwicklung eine Fülle neuer Zugangsoptionen zu standardisiert (bereits und noch in Zukunft) erhobenen Daten für den Ausbau einer dezidierten Wissensbasis zu Bildungsfragen in Ländern der Entwicklungszusammenarbeit.

Es bleibt zu beobachten, in welchem Tempo und zu welchem Ausmaß die Teilnahme von Ländern der Entwicklungszusammenarbeit an internationalen Vergleichsstudien zunehmen und sich eine globale Assessment-Kultur ausweiten wird (vgl. Smith 2016). Für die erziehungswissenschaftliche Forschung im Kontext von Bildung und Entwicklung gewinnt der Auftrag die Teilnahme, die Durchführung und insbesondere die Rezeption der Ergebnisse kritisch und analytisch zu begleiten im Zuge von *PISA for Development* eine noch größere Bedeutung. Inwiefern die OECD mit *PISA for Development* im Rahmen der *Sustainable Development Goals* zur Qualitätsentwicklung in Ländern der Entwicklungszusammenarbeit beiträgt wird sich zeigen. De facto sind bereits heute nationale, regionale und internationale Vergleiche zunehmend auch in den Diskussionen zur Bildungsqualität in Ländern der Entwicklungszusammenarbeit nicht mehr wegzudenken.

Anmerkungen

- 1 http://www.pisa.tum.de/fileadmin/w00bgi/www/Broschuere/21_Pisa-Bro_2015_Lay01_141201_Einzelseiten.pdf (Stand: 30.10.15).
- 2 Die Weltbank verwendet die World Bank Atlas Methode und kategorisiert Länder nach dem GNI (Bruttonationaleinkommen) pro Kopf (Aktualisierung Juli 2013): low income: \$1.035 oder weniger, lower middle income: \$1.036 bis \$4.085, upper middle income: 4.086 bis \$12.615, high income: \$12.616 oder mehr (<http://data.worldbank.org/news/new-country-classifications> (Stand: 01.10.15)).
- 3 Der Weltbildungsbericht schlägt als Einteilung der Länder und Regionen der Welt die folgende Gruppierung nach „EFA Regionen“ vor: Arab States (20 Länder/Regionen), Central and Eastern Europe (21 Länder), Central Asia (9 Länder), East Asia and the Pacific (33 Länder/Regionen), Latin America and the Caribbean (43 Länder/Regionen), North America and Western Europe (26 Länder/Regionen), South and West Asia (9 Länder), Sub-Saharan Africa (46 Länder), UNESCO (2015, S. 316–317).
- 4 LLECE (Laboratorio Latinoamericano de Evaluación de la Calidad de la Educación): Latin American Laboratory for Assessment of the Quality of Education (<http://www.unesco.org/new/es/santiago/education/education-assessment/>).
- 5 <http://www.sacmeq.org/?q=sacmeq-projects/sacmeq-iv> (Stand: 26.11.2015).
- 6 <http://www.education.gov.za/Curriculum/AnnualNationalAssessment/tabid/424/Default.aspx> (Stand: 06.11.15).
- 7 <http://www.iiep.unesco.org/en/our-expertise/sacmeq> (Stand: 06.11.15).
- 8 <https://efareport.wordpress.com/2014/06/03/the-challenges-and-rewards-of-measuring-global-learning-after-2015> (Stand: 27.11.2015).

Literatur

- Adair, J. G. (1995):** The research environment in developing countries. Contributions to the national development of the discipline. In: International Journal of Psychology, Vol. 30, I. 6, S. 643–662.
- Adams, R./Cresswell, J. (2016):** PISA for Development Technical Strand A: Enhancement of PISA Cognitive Instruments. OECD Education Working Papers. Paris.
- Anastasi, A. (1964):** Culture-fair testing. In: Educational Horizons, Vol. 43, I. 1, S. 26–30.
- Benavot, A./Köseleci, N. (2015):** Seeking quality in education: the growth of national learning assessments, 1990–2013; Background paper prepared for the Education for all global monitoring report 2015, Education for All 2000–2015: achievements and challenges.
- Chinapah, V. (1997):** Monitoring learning achievement. Paris.
- Döbert, H./Weishaupt, H. (2012):** Bildungsmonitoring. In: Wacker, A./Maier, U./Wissinger, J. (Hg.): Schul- und Unterrichtsreform durch ergebnisorientierte Steuerung. Empirische Befunde und forschungsmethodische Implikationen. Wiesbaden, S. 155–173 (SpringerLink Bücher).
- Hammett, D./Wedgewood, R. (Hg.) (2005):** The methodological challenges of researching education and skills development in Africa. Edinburgh.
- Jansen, S. D. (2005):** 'Lost in Translation?' Researching Education in Africa. In: Hammett, D./Wedgewood, R. (Hg.): The methodological challenges of researching education and skills development in Africa. Edinburgh, S. 16–25.
- Kellaghan, T./Greaney, V./Murray, T. S. (2009):** Using the Results of a National Assessment of Educational Achievement. Washington, DC. (National assessments of educational achievement, v. 5).
- Lang-Wojtasik, G. (2002):** Bildungsforschung im Nord-Süd-Kontext – Probleme und Forschungsimplicationen. In: Tertium Comparationis, Vol. 8, I. 2, S. 136–145.
- Lee, V. E./Zuze, T. L./Ross, K. N. (2005):** School effectiveness in 14 sub-Saharan African countries: Links with 6th Graders' reading achievement. In: Studies in Educational Evaluation, Vol. 31, I. 2–3, S. 207–246.
- Liberman, J./Clarke, M. (2012):** Review of World Bank support for student assessment strategies in client countries, 1998–2009. Washington.
- Lockheed, M. (2010):** The craft of education assessment: Does participation in international and regional assessments build assessment capacity in developing countries? An independent evaluation in IEA's Program on the Assessment of Student Achievement (PASA). Amsterdam.
- Lockheed, M. (2013):** Causes and Consequences of International Assessments in Developing Countries. In: Meyer, H.-D./Benavot, A. (Hg.): PISA, power, and policy. The emergence of global educational governance. Oxford, S. 163–184 (Oxford studies in comparative education).
- Lockheed, M. (2015):** Why do countries participate in international large-scale assessments? The case of PISA. Policy Research Working Paper 7447.

- Maier, U./Kuper, H. (2012):** Vergleichsarbeiten als Instrumente der Qualitätsentwicklung an Schule. Überblick zum Forschungsstand. In: *Die Deutsche Schule*, Vol. 104, I. 1. S. 88–99.
- Mullen, M. R. (1995):** Diagnosing Measurement Equivalence in Cross-National Research. In: *Journal of International Business Studies*, Vol. 26, I. 3, S. 573–596.
- Murimba, S. (2005):** The impact of the Southern and Eastern Africa Consortium for monitoring educational quality (SACMEQ). In: *Prospects*, Vol. XXXV, I. 1, S. 91–108.
- Naumann, J. (2005):** TIMSS, PISA, PIRLS and low educational achievement in world society. In: *Prospects*, Vol. 35, I. 2. S. 229–248.
- OECD (2013):** PISA in low and middle income countries. Education working paper no. 93. Paris.
- OECD/World Bank Group (2014):** PISA for Development. Capacity Needs Assessment: Zambia; (Online available: http://www.oecd.org/pisa/aboutpisa/Pisa%20for%20Development_Zambia_FINAL%20with%20cover.pdf, [23.11.2015]).
- Postlethwaite, T. N. (2006):** What is a 'good' cross-national study? In: Ross, K. N./Genevois, I. J. (Hg.): *Cross-national studies of the quality of education. Planning their design and managing their impact*. Paris, S. 105–120.
- Postlethwaite, T. N./Kellaghan, T. (2008):** National assessments of educational achievement.
- Prenzel, M./Drechsel, B. (2004):** „Wie PISA zeigt...“ – Konsequenzen aus internationalen Leistungsvergleichen und ihre Begründung. In: Heinze, S./Kuntze, S. (Hg.): *Beiträge zum Mathematikunterricht*. Hildesheim, S. S. 31–39.
- Rademacher, S. (2013):** Fallvergleich und Kulturvergleich. Methodologische Implikationen einer kulturvergleichenden qualitativen Forschung. In: Hummrich, M./Rademacher, S. (Hg.): *Kulturvergleich in der qualitativen Forschung. Erziehungswissenschaftliche Perspektiven und Analysen*. Wiesbaden, S. 65–78.
- Ross, K. N./Genevois, I. J. (Hg.) (2006):** *Cross-national studies of the quality of education. Planning their design and managing their impact*. Paris.
- Schwippert, K./Bos, W. (2005):** Die Daten von Vergleichsuntersuchungen und ihre Nutzungsmöglichkeiten für erziehungswissenschaftliche Forschung. In: *Tertium Comparationis*, Vol. 11, I. 2, S. 121–133.
- Smith, W. C. (2016):** The Global Testing Culture shaping education policy, perceptions, and practice (Online available: <http://www.symposium-books.co.uk/books/bookdetails.asp?bid=94> [24.01.2016]).
- Smith, M./Barrett, A. M. (2011):** Capabilities for learning to read: An investigation of social and economic effects for grade 6 learners in Southern and East Africa. In: *International Journal of Educational Development*, Vol. 31, I. 1, S. 23–36.
- UN General Assembly (2015):** Resolution adopted by the General Assembly on 25 September 2015.
- UNESCO (2015):** Education for All 2000–2015: achievements and challenges; EFA global monitoring report, 2015.
- Wedgewood, R./Hammett, D. (2005):** The loneliness of the long distance researcher: Methodological challenges of doctoral research on education and training in Africa. In: Hammett, D./Wedgewood, R. (Hg.): *The methodological challenges of researching education and skills development in Africa*. Edinburgh, S. 4–15.

Dr. Sarah Lange

ist wissenschaftliche Mitarbeiterin am Lehrstuhl für Allgemeine Pädagogik an der Otto-Friedrich-Universität Bamberg, Forschungsschwerpunkte: International und Interkulturell Vergleichende Erziehungswissenschaft, Empirische Bildungsforschung in Ländern der Entwicklungszusammenarbeit, Empirische Lehrer(fort)bildungsforschung, Unterrichtsqualitätsforschung.